

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2005, 08:40:21 ; Search time 168 Seconds

(without alignments)
433.851 Million cell updates/sec

Title: US-10-695-140-2

Perfect score: 947

Sequence: 1 APITAYSQTRGLGCIITS.....RGVAKAVDFVPVSMETMR 180

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1812044 seqs, 404927589 residues

Total number of hits satisfying chosen parameters: 1812044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA.*

- 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/prodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/prodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/prodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 17: /cgn2_6/prodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 18: /cgn2_6/prodata/1/pubpaa/US10F_PUBCOMB.pep.*
- 19: /cgn2_6/prodata/1/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/prodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	947	100.0	180	US-10-695-140-2	Sequence 2, Appli
2	947	100.0	685	US-10-974-558-2	Sequence 2, Appli
3	947	100.0	1985	US-10-259-275-42	Sequence 42, Appli
4	947	100.0	1985	US-10-639-150-2	Sequence 2, Appli
5	947	100.0	1985	US-11-006-313-42	Sequence 42, Appli
6	947	100.0	2201	US-10-029-907-3	Sequence 3, Appli
7	947	100.0	2201	US-10-309-561-3	Sequence 3, Appli
8	947	100.0	2201	US-10-789-355-3	Sequence 3, Appli
9	947	100.0	2201	US-10-686-835-3	Sequence 3, Appli
10	947	100.0	3010	US-10-467-000-1	Sequence 1, Appli
11	944	99.7	1692	US-09-919-901-4	Sequence 4, Appli

12	944	99.7	1692	14	US-10-191-966-4	Sequence 4, Appli
13	944	99.7	1985	16	US-10-492-178-1	Sequence 1, Appli
14	944	99.7	1985	16	US-10-492-178-6	Sequence 6, Appli
15	944	99.7	2307	10	US-09-919-901-2	Sequence 2, Appli
16	944	99.7	2307	14	US-10-191-966-2	Sequence 2, Appli
17	941	99.4	1692	10	US-09-919-901-11	Sequence 11, Appli
18	941	99.4	1692	10	US-09-919-901-18	Sequence 18, Appli
19	941	99.4	1692	14	US-10-191-966-11	Sequence 11, Appli
20	941	99.4	1692	14	US-10-191-966-18	Sequence 18, Appli
21	941	99.4	2307	10	US-09-919-901-9	Sequence 9, Appli
22	941	99.4	2307	10	US-09-919-901-16	Sequence 16, Appli
23	941	99.4	2307	14	US-10-191-966-9	Sequence 9, Appli
24	941	99.4	2307	14	US-10-191-966-16	Sequence 16, Appli
25	937	98.9	292	13	US-10-017-736-15	Sequence 15, Appli
26	937	98.9	292	15	US-10-650-585-15	Sequence 15, Appli
27	937	98.9	303	13	US-10-017-736-10	Sequence 10, Appli
28	937	98.9	303	13	US-10-017-736-16	Sequence 16, Appli
29	937	98.9	303	13	US-10-017-736-18	Sequence 18, Appli
30	937	98.9	303	15	US-10-650-585-10	Sequence 10, Appli
31	937	98.9	303	15	US-10-650-585-16	Sequence 16, Appli
32	937	98.9	303	15	US-10-650-585-18	Sequence 18, Appli
33	937	98.9	334	13	US-10-017-736-4	Sequence 4, Appli
34	937	98.9	334	15	US-10-650-585-4	Sequence 4, Appli
35	937	98.9	341	13	US-10-017-736-14	Sequence 14, Appli
36	937	98.9	341	15	US-10-650-585-14	Sequence 14, Appli
37	937	98.9	352	13	US-10-017-736-13	Sequence 13, Appli
38	937	98.9	352	15	US-10-650-585-13	Sequence 13, Appli
39	937	98.9	380	13	US-10-017-736-12	Sequence 12, Appli
40	937	98.9	380	15	US-10-650-585-12	Sequence 12, Appli
41	937	98.9	393	13	US-10-017-736-11	Sequence 11, Appli
42	937	98.9	393	15	US-10-650-585-11	Sequence 11, Appli
43	937	98.9	409	13	US-10-017-736-2	Sequence 2, Appli
44	937	98.9	409	15	US-10-650-585-2	Sequence 2, Appli
45	937	98.9	1736	14	US-10-328-127-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-10-695-140-2
; Sequence 2, Application US/10695140
; Publication No. US20050019753A1
; GENERAL INFORMATION:
; APPLICANT: Boehringer Ingelheim International GmbH
; TITLE OF INVENTION: Inhibitor-Resistant HCV NS3 Protease
; FILE REFERENCE: 13/10905
; CURRENT APPLICATION NUMBER: US/10/695,140
; CURRENT FILING DATE: 2003-10-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 180
; TYPE: PRT
; ORGANISM: HCV peptide
US-10-695-140-2

Query Match -100.0%; Score 947; DB 17; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.8e-92;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	APITAYSQTRGLGCIITS	LTGRDRNQVEGEVQVYVSTATQSLATCVNGCVTVYHGAG	60
Db	1	APITAYSQTRGLGCIITS	LTGRDRNQVEGEVQVYVSTATQSLATCVNGCVTVYHGAG	60
Qy	61	SKTLAGPKGPITQMTYTNVDQDLVGWQAPPGARS	LTPCTCGSSDLXVTRHADVIPVRRRG	120
Db	61	SKTLAGPKGPITQMTYTNVDQDLVGWQAPPGARS	LTPCTCGSSDLXVTRHADVIPVRRRG	120
Qy	121	DSRGSLLSRPVSYLKGSGGPGLLCPGSHAVGIFRAA	VCTRGVAKAVDFVPVSMETMR	180
Db	121	DSRGSLLSRPVSYLKGSGGPGLLCPGSHAVGIFRAA	VCTRGVAKAVDFVPVSMETMR	180

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2005, 08:37:01 ; Search time 22 Seconds
(without alignments)
610.765 Million cell updates/sec

Title: US-10-695-140-2
Perfect score: 947
Sequence: 1 APITAYSQOTRGLGCIITS.....RGVAKAVDFVPVSMETMR 180

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/aaa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/aaa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/aaa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/aaa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/aaa/PTUS_COMB.pep:*
6: /cgn2_6/prodata/1/aaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	947	100.0	1985	4	US-09-539-601-9
2	947	100.0	1985	4	US-09-539-601-12
3	947	100.0	1985	4	US-09-539-601-18
4	947	100.0	1985	4	US-09-539-601-24
5	947	100.0	2201	4	US-09-539-601-6
6	947	100.0	2201	4	US-09-539-601-15
7	947	100.0	2201	4	US-10-029-907-3
8	947	100.0	3010	4	US-09-539-601-3
9	947	100.0	3010	4	US-09-539-601-21
10	947	100.0	3010	4	US-09-539-601-27
11	944	99.7	181	3	US-09-011-961-4
12	944	99.7	631	2	US-08-936-865-1
13	944	99.7	632	3	US-09-198-723A-23
14	944	99.7	632	4	US-09-684-881-23
15	944	99.7	646	3	US-09-198-723A-69
16	944	99.7	646	3	US-09-198-723A-72
17	944	99.7	646	4	US-09-684-881-69
18	944	99.7	646	4	US-09-684-881-72
19	944	99.7	1692	3	US-09-263-933-4
20	944	99.7	1692	4	US-09-919-901-4
21	944	99.7	1692	4	US-10-191-966-4
22	944	99.7	2307	3	US-09-263-933-2
23	944	99.7	2307	4	US-09-919-901-2
24	944	99.7	2307	4	US-10-191-966-2
25	941	99.4	646	3	US-09-198-723A-66
26	941	99.4	646	4	US-09-684-881-66
27	941	99.4	1692	3	US-09-263-933-11

28	941	99.4	1692	3	US-09-263-933-18	Sequence 18, Appl
29	941	99.4	1692	4	US-09-919-901-11	Sequence 11, Appl
30	941	99.4	1692	4	US-09-919-901-18	Sequence 18, Appl
31	941	99.4	1692	4	US-10-191-966-11	Sequence 11, Appl
32	941	99.4	1692	4	US-10-191-966-18	Sequence 18, Appl
33	941	99.4	2307	3	US-09-263-933-9	Sequence 9, Appl
34	941	99.4	2307	3	US-09-263-933-16	Sequence 16, Appl
35	941	99.4	2307	4	US-09-919-901-9	Sequence 9, Appl
36	941	99.4	2307	4	US-09-919-901-16	Sequence 16, Appl
37	941	99.4	2307	4	US-10-191-966-9	Sequence 9, Appl
38	941	99.4	2307	4	US-10-191-966-16	Sequence 16, Appl
39	940	99.3	1985	4	US-09-539-601-30	Sequence 30, Appl
40	940	99.3	3010	4	US-09-539-601-33	Sequence 33, Appl
41	937	98.9	292	4	US-10-017-736C-15	Sequence 15, Appl
42	937	98.9	303	4	US-10-017-736C-10	Sequence 10, Appl
43	937	98.9	303	4	US-10-017-736C-16	Sequence 16, Appl
44	937	98.9	303	4	US-10-017-736C-18	Sequence 18, Appl
45	937	98.9	334	4	US-10-017-736C-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-539-601-9
; Sequence 9, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1985
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-539-601-9

Query Match	100.0%;	Score 947;	DB 4;	Length 1985;			
Best Local Similarity	100.0%;	Pred. No. 8.9e-89;					
Matches 180;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;			
yy	1	APITAYSQOTRGLGCIITS	LTGRDRNQVGEVQV	STATQSF	LATCNGV	CVTVYHGAG 60	
bb	2	APITAYSQOTRGLGCIITS	LTGRDRNQVGEVQV	STATQSF	LATCNGV	CVTVYHGAG 61	
yy	61	SKTLAGPKGPI	OMYTNVDQDLV	GWQAPPGARSL	TPCTCGSSD	LYLVTRHADVI	PVRRRG 120
bb	62	SKTLAGPKGPI	OMYTNVDQDLV	GWQAPPGARSL	TPCTCGSSD	LYLVTRHADVI	PVRRRG 121
yy	121	DSRGSLSPRPV	SYLKGSGG	PLLCPSGHAVG	IFRAAVCTRG	VAKAVDFVPV	ESMETTMR 180
yy	122	DSRGSLSPRPV	SYLKGSGG	PLLCPSGHAVG	IFRAAVCTRG	VAKAVDFVPV	ESMETTMR 181

RESULT 2
US-09-539-601-12
; Sequence 12, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2005, 08:34:45 ; Search time 177 Seconds
(without alignments)

520.758 Million cell updates/sec

Title: US-10-695-140-2

Perfect score: 947

Sequence: 1 APITAYSQTRGLGCIITS.....RGVAKAVDFVPVSMETTR 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	947	100.0	3010	2 Q9WMX2	Q9wmx2 hepatitis c
2	944	99.7	180	2 Q8QW30	Q8qw30 hepatitis c
3	944	99.7	3010	2 P88803	P88803 hepatitis c
4	944	99.7	3010	2 P90191	P90191 hepatitis c
5	944	99.7	3010	2 Q99AU2	Q99au2 hepatitis c
6	943	99.6	180	2 Q8GVX9	Q8gvx9 hepatitis c
7	943	99.6	181	2 Q91RP8	Q91rp8 hepatitis c
8	943	99.6	1984	2 Q7T4V8	Q7t4v8 hepatitis c
9	943	99.6	3010	2 Q9DTF0	Q9dtf0 hepatitis c
10	941	99.4	180	2 Q8QML3	Q8qml3 hepatitis c
11	941	99.4	3010	2 Q6GYR9	Q6gyr9 hepatitis c
12	941	99.4	3010	2 Q9DTP9	Q9dtp9 hepatitis c
13	940	99.3	180	2 Q8QLV3	Q8qlv3 hepatitis c
14	940	99.3	180	2 Q8QM29	Q8qm29 hepatitis c
15	940	99.3	180	2 Q8QM32	Q8qm32 hepatitis c
16	940	99.3	181	2 Q91RQ4	Q91rq4 hepatitis c
17	940	99.3	234	2 Q70823	Q70823 hepatitis c
18	940	99.3	234	2 Q70824	Q70824 hepatitis c
19	940	99.3	234	2 Q70825	Q70825 hepatitis c
20	940	99.3	234	2 Q70827	Q70827 hepatitis c
21	940	99.3	234	2 Q70829	Q70829 hepatitis c
22	940	99.3	234	2 Q70830	Q70830 hepatitis c
23	940	99.3	234	2 Q70831	Q70831 hepatitis c
24	940	99.3	234	2 Q70832	Q70832 hepatitis c
25	940	99.3	234	2 Q89253	Q89253 hepatitis c
26	940	99.3	234	2 Q89254	Q89254 hepatitis c
27	940	99.3	234	2 Q89255	Q89255 hepatitis c
28	940	99.3	234	2 Q89256	Q89256 hepatitis c
29	940	99.3	361	2 Q70818	Q70818 hepatitis c
30	940	99.3	3008	2 Q9J3F4	Q9j3f4 hepatitis c
31	940	99.3	3010	2 Q68533	Q68533 hepatitis c

32	940	99.3	3010	2 Q9J3H3	Q9j3h3 hepatitis c
33	940	99.3	3010	2 Q9QF61	Q9qp61 hepatitis c
34	940	99.3	3011	2 Q6SCJ5	Q6scj5 hepatitis c
35	939	99.2	181	2 Q91RQ5	Q91rq5 hepatitis c
36	939	99.2	3010	2 Q9DTE4	Q9dte4 hepatitis c
37	938	99.0	180	2 Q8QVY5	Q8qvys hepatitis c
38	938	99.0	180	2 Q8QW31	Q8qw31 hepatitis c
39	938	99.0	180	2 Q8QW09	Q8qwm9 hepatitis c
40	937	98.9	180	2 Q8QW08	Q8qwm8 hepatitis c
41	937	98.9	180	2 Q8QW15	Q8qwl5 hepatitis c
42	937	98.9	180	2 Q8QW22	Q8qlw2 hepatitis c
43	937	98.9	180	2 Q8QW08	Q8qwm8 hepatitis c
44	937	98.9	180	2 Q8QW11	Q8qwm1 hepatitis c
45	937	98.9	180	2 Q8QW34	Q8qwm34 hepatitis c

ALIGNMENTS

RESULT 1
Q9WMX2
ID Q9WMX2 PRELIMINARY; PRT; 3010 AA.
AC Q9WMX2;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polypeptide.
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus; Hepatitis C virus type 1.
OX NCBI_TaxID=31647;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RS;
RX MEDLINE=99322193; PubMed=10390360; DOI=10.1126/science.285.5424.110;
RA Lohmann V., Koerner F., Koch J.O., Herian U., Theilmann L.,
RA Bartenschlager R.;
RT "Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell
line";
RL Science 285:110-113(1999).
DR EMBL; AJ238799; CAB46677.1; -
DR PIR; A61196; A61196.
DR PIR; P0246; P0246.
DR PIR; PS0329; PS0329.
DR HSP; Q8YYS1; LCWX.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR001345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept_Sex_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2005, 08:35:26 ; Search time 39 Seconds
(without alignments)

Title: US-10-695-140-2

Perfect score:

Sequence: 1 APITAYSQQTRGLGCIITS.....RGVAKAVDFVPVESMETTMR 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Minimum DB seq length: 200000000
Maximum DB seq length: 200000000

Post-processing: Minimum M

Post-processing: Minimum Match 0%
Maximum Match 100%

Maximum match 100%
Listing first 45 summaries

Database : PIR 79:★

```
1: pirl:*
```

```
2: pir2:*
```

3: pir3:★

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match		Length	DB	ID	Description
		Match	Length				
1	926	97.8	3010	1	A45573	genome polyprotein	
2	923	97.5	3010	1	GNWVCJ	genome polyprotein	
3	921	97.3	3010	1	GNWVTC	genome polyprotein	
4	918	96.9	3010	1	GNWVTH	genome polyprotein	
5	904	95.5	3010	1	S18030	genome polyprotein	
6	884	93.3	3011	1	S40770	genome polyprotein	
7	882	93.1	3011	1	GNWVC3	genome polyprotein	
8	879	92.8	3011	1	GNWVCH	genome polyprotein	
9	797	84.2	3014	1	JC5620	genome polyprotein	
10	722	76.2	3033	1	GNWVJ8	genome polyprotein	
11	720	76.0	3033	1	QJ1303	genome polyprotein	
12	276	29.1	3005	2	T08841	polyprotein - dour	
13	256.5	27.1	2970	2	T08839	polyprotein - marm	
14	86.5	9.1	398	2	B71284	probable periplasm	
15	84	8.9	209	2	H83144	probable aromatic	
16	84	8.9	209	2	A84212	hypothetical prote	
17	81.5	8.6	786	2	S22155	oncogene 1 (tre-2	
18	80	8.4	233	2	T35594	hypothetical prote	
19	79.5	8.4	868	2	H81775	aconitate hydratase	
20	79	8.3	739	2	B88553	protein K04H4.2b (
21	79	8.3	3414	1	GNWVNE	genome polyprotein	
22	78.5	8.3	1236	2	E70977	hypothetical prote	
23	78	8.2	452	2	I39383	angio-associated m	
24	78	8.2	3412	1	GNWVTH	genome polyprotein	
25	76.5	8.1	495	2	B71360	hypothetical prote	
26	75.5	8.0	264	2	H38136	chymotrypsin-like	
27	75	7.9	433	2	H97199	htrA-like serine p	
28	75	7.9	1089	2	S22158	transforming prote	
29	75	7.9	1399	2	G83112	DNA-directed RNA p	

ALIGNMENTS

RESULT 1

genome polyprotein - hepatitis C virus (strain JT)
 N;Contains: capsid protein C; envelope protein M; hepatitis C virus (strain JT) (nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5)
 C;Species: hepatitis C virus
 C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
 C;Accession: A45573
 R;Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hijikata Virus Res. 23, 39-53, 1992
 A;Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier: A;Reference number: A45573; PMID:92295714; PMID:1318627

Query Match 97.8%; Score 926; DB 1; Length 3010;
Best Local Similarity 96.1%; Pred. No. 4.6e-77;

Qy	1	APITAVSQTRGLGCGIITSLTGRDRNQVEGEVQVSTATQSFATCVNGVCWTVYHGAG	60
Db	1027	APITAVAQTRGLGCGIVTSLTGRDRNQVEGEVQVSTATQSFATCVNGVCWTVFHGAG	1086
Qy	61	SKTLAPGKGPITOMYTNVDQDLGWQAPPGARSLTPTCTCGSSDLYLVTRHADVI	120
Db	1087	SKTLAPGKGPITOMYTNVDQDLGWGHAPPGARSLTPTCTCGSSDLYLVTRHADVI	1146
Qy	121	DSRGSLLSPRPVSYLKGSGGPIPLCPSPGHAVGIFRAAVCTRGVAKAVDFVPVSGMETTMR	180
Db	1147	DGPGSLLSPRPVSYLKGSGGPIPLCPSPGHAVGIFRAAVCTRGVAKAVDFVPVSGMETTMR	1206

RESULT. T. 2

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2005, 08:34:26 ; Search time 72 Seconds
(without alignments)
966.901 Million cell updates/sec

Title: US-10-695-140-2

Perfect score: 947

Sequence: 1 APITAYSQTRGLGCIITS.....RGVAKAVDPVPVSEMETTWR 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	947	100.0	180	8	Adn96855 Hepatitis
2	947	100.0	1985	5	Rao18001 Hepatitis
3	947	100.0	1985	5	Aae15729 Hepatitis
4	947	100.0	1985	5	Aae15731 Hepatitis
5	947	100.0	1985	5	Aae15717 Hepatitis
6	947	100.0	1985	5	Aae15727 Hepatitis
7	947	100.0	1985	5	Aae15728 Hepatitis
8	947	100.0	1985	5	Aae15722 Hepatitis
9	947	100.0	1985	5	Aae15730 Hepatitis
10	947	100.0	1985	8	Adj57846 HCV repli
11	947	100.0	2201	5	Abg30601 Hepatitis
12	947	100.0	2201	5	Abg30591 Hepatitis
13	947	100.0	2201	5	Abg30600 Hepatitis
14	947	100.0	2201	5	Abg30581 Hepatitis
15	947	100.0	2201	5	Abg30593 Hepatitis
16	947	100.0	2201	5	Abg30582 Hepatitis
17	947	100.0	2201	5	Abg30580 Hepatitis
18	947	100.0	2201	5	Abg30587 Hepatitis
19	947	100.0	2201	5	Abg30599 Hepatitis
20	947	100.0	2201	5	Abg30594 Hepatitis
21	947	100.0	2201	5	Abg30598 Hepatitis
22	947	100.0	2201	5	Abg30595 Hepatitis
23	947	100.0	3010	5	Abg32458 Hepatitis
24	947	100.0	3010	5	Abg32459 Hepatitis
25	947	100.0	3010	5	Abg32451 Hepatitis

ALIGNMENTS

RESULT 1

ADN96855
ID ADN96855 standard; protein; 180 AA.

XX ADN96855;

XX AC AC (first entry)

DT 29-JUL-2004 (first entry)

XX Hepatitis C virus NS3 protease.

XX NS3 protease; mutation; protease inhibitor; enzyme.

XX Hepatitis C virus.

XX WO2004039970-A1.

XX 13-MAY-2004.

XX 24-OCT-2003; 2003WO-CA001636.

XX 29-OCT-2002; 2002US-0421943P.

XX (BOEH) BOEHRINGER INGELHEIM INT GMBH.

XX Kukolj G, Lagace L, Marquis M, Lamarre D, Pause A;

XX WPI; 2004-400164/37.

XX N-PSDB; ADN96854.

XX Novel hepatitis C virus non-structural 3 protease, useful for identifying

XX potential second generation inhibitor of hepatitis C virus non-structural

XX 3 protease.

XX Claim 5; SEQ ID NO 2; 40pp; English.

XX The invention relates to a hepatitis C virus (HCV) NS3 protease (I), or an amino acid sequence that is 90% identical to (I), comprising one or more mutations in at least one amino acid position Arg155Gln, Arg155Trp, Ala156Gly, Ala156Thr, Ala156Val. (I) is useful for identifying a potential second generation inhibitor of HCV NS3 protease activity which involves incubating an inhibitor-resistant (I) in the presence or absence of a candidate second generation inhibitor compound, and measuring the protease activity of the inhibitor-resistant (I) in the presence and absence of the candidate second generation inhibitor compound, where a decrease in the activity of the protease in the presence of a candidate second generation inhibitor indicates that the compound inhibits the inhibitor-resistant (I). Vectors (II) containing the gene encoding (I) are useful for evaluating HCV NS3 protease activity of inhibitor-

Abg32455 Hepatitis
Abg32457 Hepatitis
Abg32460 Hepatitis
Abg32461 Hepatitis
Abg32454 Hepatitis
Abg32456 Hepatitis
Aaw13769 Hepatitis
Aaw56479 Amino aci
Aaw37807 Nonstruct
Aay17897 Native HC
Aay17894 HCV NS4A-
Abu09574 HCV Met-N
Adr38450 Hepatitis
Abg30586 Hepatitis
Abg30589 Hepatitis
Abg30583 Hepatitis
Aay70064 Recombina
Aae15720 Hepatitis